

From Images to Knowledge with ImageJ & Friends

virtual conference

Nov 30 - Dec 2, 2020

Stephan Preibisch, Stephan Saalfeld, Anna Kreshuk,
Pavel Tomancak and Virginie Uhlmann

hhmi | janelia
Conferences



Harnessing BIAFLOWS, a web platform to reproducibly deploy and benchmark Bioimage Analysis Workflows

Tutors: Sébastien Tosi (sebastien.tosi@irbbarcelona.org)
Raphaël Marée (raphael.maree@uliege.be)
Volker Baecker (volker.baecker@mri.cnrs.fr)
Romain Mormont (r.mormont@uliege.be)
Benjamin Pavie (benjamin.pavie@kuleuven.vib.be)

Session 1: 2020-12-01 08:00 UTC – 2020-12-01 12:00 UTC

Session 2: 2020-12-02 12:00 UTC – 2020-12-02 16:00 UTC

Harnessing BIAFLOWS, a web platform to reproducibly deploy and benchmark Bioimage Analysis Workflows

Keywords: reproducibility, image analysis, benchmarking, cloud computing, image database, Docker, GitHub, BIAFLOWS

Abstract

Image analysis is key to extract quantitative information from scientific microscopy images. In this tutorial we explore how image analysis workflows can be deployed in a reproducible way. Reproducibility is at the core of the scientific method itself. If a research result can not be reproduced the conclusions implied are not considered valid. This means that the image analysis that is part of a scientific project must also be reproducible.

It might seem that in the deterministic world of computation reproducibility could be easily achieved. However in practice it turns out that the reproducible sharing of image analysis workflows is a complex task. Some key points for achieving reproducibility in this context are:

- the availability of the original data
- the availability of the original metadata, including the parameters of the workflow at hand
- the availability of the original source code of the workflow
- a clear path of how to get an executable from the source code
- the availability of the original computational environment, including the operating system, the software platform, libraries and plugins
- the documentation of how to use the workflow
- an access to the computational resources needed to execute the workflow

In the context of the COST Action (CA15124) NEUBIAS (Network of European BioImage Analysts), we developed a cloud based user friendly web platform BIAFLOWS, that helps to make image analysis reproducible. In this tutorial, participants will learn how to use BIAFLOWS to run and compare existing workflows and how to integrate their own workflows and data into BIAFLOWS.

Description

The purpose of this tutorial is to understand the problems and solutions related to reproducible image analysis workflows and benchmarking. It will teach the participants how to use BIAFLOWS and should provide a good starting point to developers and bio-image-analysts interested in integrating their own image analysis workflows and annotated datasets into BIAFLOWS. The plan that will be followed (subject to minor updates):

1. The problem of reproducible image analysis and benchmarking (15 minutes)
 - 1.1. Reproducible image analysis
 - 1.2. Benchmarking
2. Quick round table of the participants (20 minutes)
Focus on background, experience and needs and interest for the platform.
3. Introduction to BIAFLOWS (20 minutes)
 - 3.1. Presentation
 - 3.2. Demo
4. Adding content to BIAFLOWS (3h)
 - 4.1. Uploading images to the sandbox server
 - 4.2. Running a workflow on existing images and retrieving benchmark results
 - 4.3. Adding a new workflow by using an existing workflow as template
 - 4.4. Introduction on how to set up a complete development environment to compile Docker images, compile and test containers locally and interact with BIAFLOWS

Target: Bioimage analysis workflows developers and bioimage analysts

Maximum number of participants: 10

Pre-requirements:

- A personal computer with Docker installed <https://docs.docker.com/>
- Familiarity with Python programming language (or at least a programming language)
- A Python 3.6 environment installed (<https://www.anaconda.com/> recommended)
- Familiarity with ImageJ and ImageJ macro language
- Prepare an ImageJ macro to process images from one of these two projects:
<https://biaflows.neubias.org/#/project/5955/images> (easy task)
<https://biaflows.neubias.org/#/project/12182234/images> (challenging task)
Click *Try Online* when opening these links. The images can then be downloaded by clicking on the blue arrows on the left side. The macro should return one label mask where each nucleus holds a unique nonzero ID, and it should be able to process all the images from an input folder and save the results as TIFF images to an output folder.

Tutors (all tutors available from December 1st to Dec 2nd from 9 a.m. to 5 p.m. CET)

Volker Baecker (Montpellier Ressources Imagerie, Biocampus Montpellier)

Volker Baecker completed a MSc degree in Computer Science at the University of Dortmund in 2000. He then worked as a software engineer at Georg Heege for 3 years. Since 2004 he has been a bioimage analyst at the core imaging facility Montpellier Ressources Imagerie (CNRS, INSERM, UM). In the context of the EU COST Action NEUBIAS, he is involved in the development of the Bioimage Informatics Index and the BioImage Analysis workflows benchmarking platform BIAFLOWS.

Benjamin Pavie (VIB BioImaging Core)

Benjamin Pavie completed a MSc degree in Fundamental Virology at Pasteur Institute Paris/Paris VI and a bachelor in Java and Database. He then worked as software developer for 5 years at Scito and then 6 years as computational biologist at UT Southwestern, Dallas, USA working on image analysis on cell heterogeneity in the Altschuler and Wu Lab. He moved to Belgium to work for the VIB as Bio-Image Analyst at the Microscopy Core facility developing tools and helping members of the institute / university in their microscopy image analysis. He is also involved in the development of BIAFLOWS for reproducible benchmarking in the context of the EU COST NEUBIAS action.

Romain Mormont (Institute Montefiore, University of Liège)

Romain Mormont completed a MSc degree in Computer Science and Engineering (with focus on Intelligent Systems) at the University of Liège in 2016. During his master thesis, he studied the automatic segmentation, extraction and classification of biological features from whole-slide images containing thyroid nodule cell samples. Since then, he has been a PhD student at ULiège. His work focuses mostly on developing and studying machine learning algorithms that can cope with "small-data" problems. He has investigated transfer and multi-task learning approaches to tackle such problems. He has a strong experience in machine/deep learning, programming (web, python, java...), and digital pathology. He is also core developer of BIAFLOWS for reproducible benchmarking in the context of the EU COST NEUBIAS action.

Raphaël Marée (Institute Montefiore, University of Liège)

Raphaël Marée obtained his MSc degree in computer science at ULiège in 2000 and the PhD degree in computer science in 2005. During his PhD thesis, he proposed generic machine learning algorithms for image classification that were later extended with colleagues for content-based image retrieval, landmark detection, and pixel-level segmentation. After his PhD thesis, he was manager of the GIGA Bioinformatics core facility (2005-2014) where he provided data analysis services for biomedical researchers. He spent one year (2014-2015) as postdoctoral researcher at Institut Pasteur (Paris, France) to work on the development of image analysis tools for nephrology and digital pathology. In 2010, he initiated the Cytomine research project and he has been the Cytomine-ULiège scientific coordinator since then. His current research interests are in the broad area of machine/deep learning techniques, web software development methodologies, and open science paradigms, with specific focus on their applications to various problems in biology and medicine that involve very large sets of high-resolution images. He is also currently member of the management committee of the EU COST NEUBIAS and

COMULIS networks, and co-lead of their software development work packages. He has also been involved into the organization of several scientific events related to computer vision and biomedical imaging (e.g. 1st and 2nd Computational Pathology ECCB workshop 2016-2018). He has been reviewer for many scientific journals (IEEE Transactions, Nature Methods, Bioinformatics,...), biomedical imaging conferences (MICCAI, ISBI,...) and international funding agencies (H2020 Marie Curie COFUND and ERC, UK Medical Research Council). Raphaël Marée has published over 80 papers (h-index 23).

Sébastien Tosi (Advanced Digital Microscopy Core Facility, IRB Barcelona)

Sébastien Tosi completed a MSc degree in Optics and Signal Processing (Centrale Marseille, France) and a PhD in Data Storage (University of Limerick, Ireland). He then worked as a system engineer for 3 years at SIDSA, a Telecommunication company (Madrid, Spain). He is now at Advanced Digital Microscopy Core Facility of IRB Barcelona (Barcelona, Spain) since 2010 where he works as a bioimage analyst and research engineer handling and designing microscopes. He has been organizing several international training schools (EMBL Master Course for BioImage Data Analysis 2014-2017, University of Copenhagen Image analysis PhD Course 2019-2020), authored numerous scientific articles and book chapters on bioimage analysis and led NEUBIAS Workgroup 5 to develop BIAFLOWS (2016-2020).